Математика и механика

УДК 519.872:681.324

А. Н. ДУДИН, ХАЛАФ ИХСАН

НАХОЖДЕНИЕ ОПТИМАЛЬНОЙ ГИСТЕРЕЗИСНОЙ СТРАТЕГИИ УПРАВЛЕНИЯ ДВУХСКОРОСТНОЙ СИСТЕМОЙ МАССОВОГО ОБСЛУЖИВАНИЯ

Многие задачи оптимизации функционирования реальных систем, такие, например, как задача нахождения оптимальной стратегии запроса канального ресурса в сетях связи с предоставлением каналов по требованию [1], задача нахождения оптимальной текущей границы между информационными зонами коммутации пакетов и коммутации каналов низкого приоритета при использовании принципов гибридной коммутации с плавающим порогом или адаптивной коммутации в цифровых сетях интегрального обслуживания [2] и т. д., могут быть успешно решены с помощью следующей математической модели оптимизации функционирования системы массового обслуживания (СМО) с управляемой скоростью обслуживания.

На вход однолинейной СМО с ожиданием поступает стационарный пуассоновский поток заявок интенсивности λ. Прибор может работать в двух режимах. При работе прибора в l-ом режиме время обслуживания заявки имеет распределение $B_l(t)$, $l=\overline{1,2}$. Изменение режима работы прибора возможно в моменты окончания обслуживания заявок. Качество функционирования системы оценивается функционалом потерь в единицу времени вида:

$$I=aV+c_1P_1+c_2P_2+dM, \qquad (1)$$

где V — среднее время пребывания заявки в системе; a — стоимость единицы времени пребывания заявки в системе; P_l — средняя доля времени работы прибора в l-ом режиме; c_l — стоимость единицы времени работы в l-ом режиме, $l=\overline{1,2}$; M — среднее число переключений режима работы прибора в единицу времени; d — плата за одно такое переключение.

Необходимо указать оптимальную стратегию выбора режима работы прибора. Эта задача является нетривиальной, если выполняется следующее. Пусть режимы работы прибора занумерованы так, что $b_{11}>b_{12}$, где $b_{1l}=\int\limits_0^\infty(1-B_l(t))\,dt$ — среднее время обслуживания заявки в l-ом режиме, $l=1,2$. Тогда требуется, чтобы выполнялись неравенства:

$$a>0,\quad c_1\leqslant c_2,\quad d\geqslant 0,\quad b_{12}<\lambda^{-1}. \qquad (2)$$

В случае, когда $d=0$, т. е. плата за переключение отсутствует, известно, что стратегия, оптимальная в классе всех марковских стратегий, является пороговой [3]. В случае $d>0$ рекомендуется искать оптимальную стратегию в классе гистерезисных стратегий, т. е. стратегий следующего вида. Фиксируются натуральные числа j и k, называемые порогами, $0\leqslant j\leqslant k<+\infty$. Если число i заявок в системе в момент окончания обслуживания заявки удовлетворяет неравенству $i\leqslant j$, то следующая заявка обслуживается в первом режиме. Если $i>k$, то следующая заявка обслуживается во втором режиме. Если же $j<i\leqslant k$, то следующая заявка обслуживается в том же режиме, что и предыдущая. Для нахождения оптимальной гистерезисной стратегии, таким образом, надо оптимальным образом выбрать пороги j и k. В случае произвольных распределений $B_l(t)$, $l=\overline{1,2}$, задача нахождения оптимальных порогов не решена. Приведем ее решение. Для решения применяется так называемый прямой подход [4]. Зафиксируем значения порогов j и k. Найдем стационарное распределение вероятностей состояний системы (числа заявок в системе) в моменты окончания обслуживания заявок. Используя его, получаем явную зависимость значений величин V, P_1, P_2, M, входящих в функционал (1), от j и k. Тем самым задача поиска оптимального набора порогов j, k будет сведена к решению задачи минимизации функции двух переменных.

Несложно убедиться, что двумерный случайный процесс $\{i(t), \nu(t)\}$, где $i(t)$ — число заявок в системе в момент t, а $\nu(t)$ — номер режима, который использовался перед моментом t, заданный для $t\in T$, где T — множество моментов окончания обслуживания заявок, является двумерной цепью Маркова. При выполнении последнего из условий (2) существуют стационарные распределения вероятностей: $\pi_i=\lim\limits_{t\to\infty} P\{i(t)=i,\ \nu(t)=1\}$, $i\geqslant 0$, $\varkappa_i=\lim\limits_{t\to\infty} P\{i(t)=i,\ \nu(t)=2\}$, $i\geqslant j$. Введем производящие функции:

$$\Pi_1(z)=\sum_{i=0}^{j-1}\pi_i z^i,\quad \Pi_2(z)=\sum_{i=j}^{k}\pi_i z^i,\quad \Pi_3(z)=\sum_{i=k+1}^{\infty}\pi_i z^i,$$

$$K(z)=\sum_{i=j}^{\infty}\varkappa_i z^i,\ |z|\leqslant 1.$$

Теорема 1. Производящие функции $\Pi_l(z)$, $l=1, 3$, $K(z)$ задаются следующими соотношениями:

$$\Pi_1(z)=\sum_{i=0}^{j-1}\alpha_i^{(1)} z^i\pi_0, \tag{3}$$

$$\Pi_2(z)=z^j\left[\sum_{l=0}^{k-j}\alpha_{j+l}^{(1)}z^l-v(k,\ j)\sum_{l=0}^{k-j}\Delta_l z^l\right]\pi_0, \tag{4}$$

$$\begin{gathered}\Pi_3(z)=z^{-1}\{(\beta_1(\lambda(1-z))-z)(\Pi_1(z)+\Pi_2(z))-\\ -\beta_1(\lambda(1-z))[(1-z)-z^j v(k,\ j)]\pi_0\},\end{gathered} \tag{5}$$

$$\begin{gathered}K(z)\frac{z^{-1}\beta_2(\lambda(1-z))}{z-\beta_2(\lambda(1-z))}[(\beta_1(\lambda(1-z))-z)(\Pi_1(z)+\Pi_2(z))-\\ -\pi_0(1-z)\beta_1(\lambda(1-z))+z^j v(k,\ j)(\beta_1(\lambda(1-z))-z)\pi_0],\end{gathered} \tag{6}$$

$$\pi_0=(1-\rho_2)\left\{1+(\rho_1-\rho_2)\left[\sum_{i=0}^{k}\alpha_i^{(1)}-v(k,\ j)\left(\sum_{i=0}^{k-j}\Delta_i-1\right)\right]\right\}^{-1}, \tag{7}$$

где $\alpha_i^{(1)}$ — коэффициент при π_0 в выражении π_i у классической СМО $M|G|1$ с функцией распределения времени обслуживания $B_1(t)$,

$$\Delta_i=\sum_{l=0}^{i}\alpha_l^{(1)},\ i\geqslant 0,\ \rho_l=\lambda b_{1l},\ \beta_l(s)=\int_0^\infty e^{-st}dB_l(t),\ l=\overline{1,2};\ v(k,j)=\frac{\alpha_{k+1}^{(1)}}{\Delta_{k-j+1}}.$$

Доказательство. Используя формулу полной вероятности, можно убедиться, что стационарные вероятности π_i, $\varkappa_i$ удовлетворяют следующей системе линейных алгебраических уравнений:

$$\pi_i=\pi_0 f_i^{(1)}+\sum_{l=1}^{i+1}\pi_l f_{i-l+1}^{(1)},\ i=0,\ \overline{j-2}, \tag{8}$$

$$\pi_i = \pi_0 f_i^{(1)} + \sum_{l=1}^{i+1} \pi_l f_{i-l+1}^{(1)} + \varkappa_j f_{i-j+1}^{(1)}, \quad i = \overline{j-1,\ k-1}, \tag{9}$$

$$\pi_i = \pi_0 f_i^{(1)} + \sum_{l=1}^{k} \pi_l f_{i-l+1}^{(1)} + \varkappa_j f_{i-j+1}^{(1)}, \quad i \geqslant k, \tag{10}$$

$$\varkappa_i = \sum_{l=j+1}^{i+1} \varkappa_l f_{i-l+1}^{(2)} + \sum_{l=k+1}^{i+1} \pi_l f_{i-l+1}^{(2)}, \quad i \geqslant j, \tag{11}$$

где $f_m^{(l)} = \int\limits_0^\infty \frac{(\lambda t)^m}{m!} e^{-\lambda t} dB_l(t)$ — вероятность поступления m заявок за время обслуживания заявки в l-ом режиме, $m \geqslant 0$, $l = \overline{1,\ 2}$. Решая систему (8)—(11) с учетом условия нормировки, убеждаемся в справедливости (3)—(7).

Справедлива следующая

Теорема 2. Явная зависимость значения функционала (1) от порогов j, k задается следующей формулой:

$$I = \frac{Q + Rw^{(0)}(k,\ j) + \tilde{a}(\rho_1 - \rho_2)\left[w^{(1)}(k,\ j) - j\left(w^{(0)}(k,\ j) - w^{(0)}(k,\ k)\right)\right] + 2\lambda d(1 - \rho_2)\, v(k,\ j)}{1 + (\rho_1 - \rho_2)\, w^{(0)}(k,\ j)}, \tag{12}$$

где $w^{(m)}(k,\ j) = \sum\limits_{i=0}^{k} i^m \alpha_i^{(1)} - v(k,\ j) \sum\limits_{i=1}^{k-j} i^m \Delta_i$, $m = \overline{0,\ 1}$,

$$Q = \tilde{a}\left(\rho_1 + \delta_2^2 (1 - \rho_2)^{-1}\right) + c_2\rho_2,$$

$$R = \frac{\tilde{a}\,\delta_1^2 (1 - \rho_2) - \delta_2^2 (1 - \rho_1)}{1 - \rho_2} + c_1 \rho_1 (1 - \rho_2) - c_2\rho_2 (1 - \rho_1), \tag{13}$$

$$\delta_l^2 = \lambda^2 b_{2l}/2, \quad b_{2l} = \int\limits_0^\infty t^2 d\,B_l(t), \quad l = \overline{1,\ 2}, \quad \tilde{a} = a\lambda^{-1}.$$

Доказательство. Можно убедиться, что для данной системы справедлива формула Литтла: $\lambda V = L$, где L — среднее число заявок в системе в моменты окончания обслуживания заявок. А величина L, очевидно, может быть подсчитана по формуле

$$L = \frac{d}{dz}\left(\sum_{l=1}^{3} \Pi_l(z) + K(z)\right)\Big|_{z=1}.$$

Далее, используя закон больших чисел, можно убедиться, что $P_1 = \rho_1 \sum\limits_{l=1}^{3} \Pi_l(1)$, $P_2 = \rho_2 K(1)$.

Очевидно, что переключение со второго режима на первый происходит в момент попадания процесса $\{i(t),\ \nu(t)\}$ в состояние $(j,\ 2)$. Поэтому вероятность того, что в данный момент окончания обслуживания заявки произойдет такое переключение, равна $\varkappa_j$. Учитывая, что среднее число моментов окончания обслуживания заявок в единицу времени равно λ, а также, что среднее число переключений со второго режима на первый равно среднему числу переключений с первого режима на второй, имеем: $M = 2\lambda\varkappa_j = 2\lambda v(k,\ j)\pi_0$. Подставляя полученные выражения для V, P_1, P_2 и M в (1), после преобразований получаем (12)—(13). Теорема 2 доказана.

Следствие. Если распределения $B_l(t)$ — экспоненциальные, то соотношение (12) еще более упрощается: $I = \dfrac{\tau(1 - \rho_1^{k-j+2}) + \rho_1^{k+1} \sum\limits_{0 \leqslant l+m \leqslant 2} \delta_{l,\,m} j^l k^m}{\varkappa(1 - \rho_1^{k-j+2}) + \rho_1^{k+1} \sum\limits_{0 \leqslant l+m \leqslant 1} \xi_{l,\,m} j^l k^m}$, если $\rho_1 \neq 1$, и $I = \dfrac{\sum\limits_{0 \leqslant l+m \leqslant 3} \Theta_{l,m} j^l k^m}{\sum\limits_{0 \leqslant l+m \leqslant 2} \omega_{l,m} j^l k^m}$, если $\rho_1 = 1$, где τ, $\varkappa$, $\xi_{l,m}$, $\delta_{l,m}$, $\Theta_{k,m}$, $\omega_{l,m}$ — константы, зависящие от параметров системы, вид которых не приводится из соображений экономии места.

Имея соотношения (12), (13), несложно реализовать на ЭВМ процедуру минимизации функции двух переменных и получить в итоге оптимальные параметры (j^*, k^*) гистерезисной стратегии управления режимом работы данной СМО.

Замечание 1. Имея формулы (3)—(7), можно получить явную зависимость от j, k функционала качества функционирования системы и при более сложном, чем (1), его виде.

Замечание 2. Результат теоремы 1 остается справедливым, если допустить, что с изменением режима работы прибора меняется и интенсивность поступающего потока. Надо только в аргументе преобразований Лапласа — Стилтьеса $\beta_\nu(\lambda(1-z))$ поставить у λ индекс ν, $\nu = 1,2$.

Замечание 3. Из замечания 2 вытекает, что результаты исследования данной модели можно применять и для решения различных задач динамического управления потоками в сетях связи.

Список литературы

1. Гуд В. // ТИИЭР. 1984. Т. 72. № 11. С. 180.
2. Самойленко С. И. // Методические материалы и документация по пакетам прикладных программ. 1983. Ч. 1. Вып. 24. С. 43.
3. Crabill T., Gross P., Magazine M. // Oper. Res. 1977. V. 25. № 2. P. 219.
4. Горцев А. М., Назаров А. А., Терпугов А. Ф. Управление и адаптация в системах массового обслуживания. Томск, 1978.

Поступила в редакцию 22.03.88.

УДК 519:612

В. А. БАСИК

ОБ ОДНОМ НЕЯВНОМ АЛГОРИТМЕ РЕШЕНИЯ ГРАНИЧНЫХ ЗАДАЧ ДЛЯ РАЗНОСТНЫХ ТРЕХТОЧЕЧНЫХ УРАВНЕНИЙ

При решении задачи Дирихле для дифференциальных уравнений второго порядка эллиптического типа по методу сеток возникают системы линейных алгебраических уравнений вида [1—3]:

$$A_i V_{i-1} - B_i V_i + C_i V_{i+1} = F_i, \quad i = 1, 2, \ldots, M-1, \tag{1}$$

$$V_0 = F_0, \quad V_M = F_M, \tag{2}$$

где A_i, C_i — диагональные квадратные матрицы порядка N с положительными элементами на главной диагонали; B_i — квадратные трехдиагональные матрицы того же порядка; F_i — известные векторы. Систему вида (1)—(2) называют [1] также граничной задачей для векторных трехточечных уравнений.

Использование для решения задачи (1)—(2) метода матричной прогонки требует [1] выполнения $\approx CN^3M$ арифметических операций и запоминания $\approx N^2M$ чисел — результатов промежуточных вычислений. Методы полной редукции, быстрого преобразования Фурье требуют значительно меньшего объема вычислений [1], но применимы только к частным видам уравнения (1). Ниже строится итерационный алгоритм решения задачи (1)—(2), в основу которого положена комбинация идей модифицированного метода матричной прогонки [1] и неявных итерационных методов, предложенных В. П. Ильиным в [4] и [5]. При построении алгоритма используется близость матриц A_i, B_i, C_i к матрицам A_{i+1},